Chapter 1

Why: Biology By the Numbers

"Make things as simple as possible, but not simpler." A. Einstein

Chapter Overview: In Which the Notion of Models in Biology is Examined

Charles Darwin once said "All observation should be for or against some view if it is to be of any service," meaning that data is really most meaningful against the backdrop of some conceptual or theoretical framework. The goal of this chapter is to develop a sense of the kinds of models that must be put forth to greet modern biological (quantitative) data. We will argue that any good model has to overlook some of the full complexity and detail of a given biological problem in order to generate an abstraction that is simple enough to be easily grasped by the human mind, as an aid to developing intuition and insight. At the same time, it is critical that useful models make reasonably accurate predictions, so they must include at least some of the realistic details of the biological system. The art of model building lies in striking the proper balance between not enough detail and too much. A second thread of the chapter centers on the role of having a feeling for the numbers: sizes, shapes, times and energies associated with biological processes. Here we introduce the style of making numerical estimates that will be used throughout the book.

1.1 Physical Biology of the Cell

With increasing regularity, the experimental methods used to query living systems and the data they produce are quantitative. For example, measurements of gene expression in living cells report how much of a given gene product there is at a given place at a given time. Measurements on signaling pathways examine the extent of a cellular response such as growth, DNA replication, or actin

polymerization, as a function of the concentration of some upstream signaling molecule. Other measurements report the tendency of DNA to wrap into compact structures known as nucleosomes as a function of the underlying sequence. The list of examples goes on and on and is revealed by the many graphs depicting quantitative measurements on biological problems that grace the current research literature.

The overarching theme of the developing field of physical biology of the cell is that quantitative data demands quantitative models and conversely, that quantitative models need to provide experimentally testable quantitative predictions about biological phenomena. What this means precisely is that in those cases where a biological problem has resulted in a quantitative measurement of how a particular biological output parameter varies as some input is changed (either because it is manipulated by the experimenter, or because it varies as part of a natural biological process), the interpretation of that biological data should also be quantitative. In the chapters that follow, our aim is to provide a series of case studies involving some of the key players and processes in molecular and cell biology that demonstrate how physical model building can respond to this new era of quantitative data and sharpen the interplay between theory and experiment.

Model Building Requires a Substrate of Biological Facts and Physical (or Chemical) Principles

One of the first steps in building models of quantitative experiments like those described above is trying to decide which features of the problem are central and which are not. However, even before this, we require facts about our system of interest. Throughout this book, we will assert a series of "facts" which fall into several different categories of certainty. We are most comfortable with statements describing observations which you as an individual observer can readily confirm, such as the statement that cells contain protein molecules. A second layer of facts with which we are nearly as comfortable are those arrived at and repeatedly confirmed by decades of experimentation using many different, independent kinds of experimental techniques. One fact of this kind is the statement that protein molecules within cells are synthesized by a large piece of machinery called the ribosome. A third category of "facts" which we will take care to mark as speculative are those which stand as compelling explanations for biological observations, but which may rest on unproven or controversial assumptions. In this category we would include the commonly accepted proposition that modern day mitochondria are the descendants of what was once a free living bacterium that entered into a symbiotic relationship with an ancestral eukaryotic cell. In these cases we will generally explain the observation and the interpretation that underlies our modeling approach. One of the most basic classes of facts that we will need throughout the remainder of the book is the nature of the molecules that make up cells and organisms. Because this will set up the common molecular language for the rest of our discussions, we will describe these molecules at some length in the next section.

Complementary to these biological observations, we embrace the proposition

that biological entities cannot violate the laws of physics and chemistry. Many fundamental physical and chemical principles can be directly applied to understanding the behavior of biological systems, for example, the physical principle that the average energy of a molecule increases with increasing environmental temperature, and the chemical principle that oil and water do not mix. Like the biological facts described above, such principles will serve as the basis for the models to be described throughout the book.

1.2 The Stuff of Life

Scientists and other curious humans observe the natural world around them and try to make sense of its constituent parts and the ways that they interact. One particular subset of the natural world called life has always held particular fascination. Any three year old child knows that a rock is not alive but that a puppy, a tree, and indeed, the child himself, are alive. Yet the three year old would be hard pressed to produce a rigorous definition of life and indeed, scientists who have thought about this for decades can only approximate a description of this fascinating property. No single feature can reliably discriminate life from nonlife, but all living systems do share certain central properties. Living systems grow, consume energy from their environment, reproduce to form offspring that are rather similar to themselves, and die.

In the quest to understand the nature of life, scientists over hundreds of years have investigated the properties of living organisms and their constituent materials with an overarching goal of trying to understand why a puppy is different from a rock, even though they are both made of atoms that obey the same physical laws. Although this quest is still ongoing, some useful generalizations about the material nature of life are now widely accepted. For example, one observed feature of many of the molecules that make up living organisms is that they tend to be relatively large and structurally complex, and are hence called macromolecules. Living organisms also contain a large number of small, simple molecules which are critical to their function ranging from water and metal ions to sugars such as glucose. Although these small molecules are important to life processes, and indeed, are used by the cell as building blocks to make the macromolecules, the distinction is usually drawn between small molecules which can arise through nonliving chemical processes and characteristic biological macromolecules which are found nowhere but in living organisms. Despite the diversity of the kinds of macromolecules found in living organisms, the kinds of atoms found within them is surprisingly restricted, consisting mainly of carbon (C), oxygen (O), nitrogen (N) and hydrogen (H) with a smattering of sulfur (S) and phosphorus (P). It is largely the ways in which this humble suite of atoms are connected to each other that creates the special properties of the macromolecules found in living systems. Of course, many other kinds of atoms and ions are critical for the biochemical processes of life, but they are generally not covalently incorporated in macromolecular structures.

Organisms Are Constructed From Four Great Classes of Macromolecules

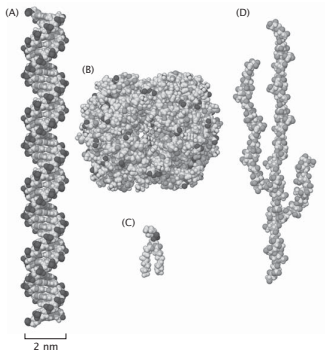

Figure 1.1: Atomic-level structural representation of members of each of the major classes of macromolecules, all drawn at the same scale. (A) Atomic structure of a small fragment of the nucleic acid DNA in the B-form, (B) atomic structure of the oxygen-carrying protein hemoglobin, (C) phosphatidylcholine lipid molecule from a cell membrane, (D) branched complex carbohydrate (M41 capsular polysaccharide) from the surface of the bacterium *E. coli*. (Illustrations courtesy of David Goodsell)

If we take any organism ranging from a small soil bacterium to a giant redwood tree or a massive whale and separate out its constituent macromolecules, we will find a small number of distinct classes, chief among these being proteins, carbohydrates, lipids and nucleic acids. Fig. 1.1 shows representative examples of each of these classes. Broadly speaking, these different classes of macromolecules perform different functions in living organisms. For example, proteins have a range of functions including that they form structural elements within cells and are largely responsible for catalyzing specific chemical reactions necessary to life. Lipids form membrane barriers that separate the cell from the outside world and subdivide the cell into interior compartments called organelles. Carbohydrates are used for energy storage, for creating specific surface properties on the outside of cell membranes and sometimes, for building rigid structural units such as cell walls. Nucleic acids perform a critical function as the memory and operating instructions that enable cells to generate all the other kinds of macromolecules and to replicate themselves. Within each of these broad categories there is tremendous diversity. A single organism may have tens of thousands of structurally distinct proteins within it. Similarly, the membranes of cells can comprise hundreds of different lipid species.

Why are these particular classes of macromolecules the stuff of life? A complete understanding of the answer to this question would require a knowledge of the specific chemical and physical conditions on the early Earth that gave rise to the first living cells, which are still active areas of research. However,

by studying living organisms on Earth today, it is seen that these molecules are wonderfully suited for the perpetuation of life for several reasons. First, each of these classes of molecules can be assembled by the cell from a small number of simpler subunit or precursor molecules. It is combinatorial assembly of these simple subunits that gives rise to the tremendous structural diversity mentioned above. What this means is that a cell needs only a relatively restricted repertoire of chemical reactions to be able to synthesize these sets of subunits from the food in its environment. The fact that most cells on Earth make a living by consuming other living organisms is probably both a cause and a consequence of the fact that they all share very similar suites of these fundamental building blocks.

Nucleic Acids and Proteins Are Polymer Languages With Different Alphabets

The rest of this book will be largely devoted to exploring the special properties of biological macromolecules and the ways that they are used by cells with the large scale goal of understanding how the special properties of life may emerge from fundamental physical and chemical interactions. We will begin by discussing one specific and fundamentally important aspect of macromolecular behavior that is necessary for life as we know it. In particular, a great triumph of the study of molecular biology in the middle part of the twentieth century was the profound realization that the sequence of nucleic acid subunits in the cell's DNA is directly responsible for determining the sequence of amino acids in that same cell's proteins (i.e. the genetic code). The relationship between DNA sequence and protein sequence has been described by Francis Crick as comprising the "two great polymer languages" of cells. As shown in fig. 1.2, both nucleic acids and proteins are built up from a limited alphabet of units.

The nucleic acids are built up from an alphabet of four letters that each represent a chemically distinct "nucleotide". (A - Adenine, G - Guanine, T - Thymine and C - Cytosine, for DNA). The chemical structure of these four nucleotides is shown in fig. 1.3. The protein language is constructed from an alphabet of 20 distinct amino acids. How can just four nucleotide bases be translated into protein sequences containing 20 different amino acids? The key realization is that the DNA letters are effectively read as "words", called codons, made up of three sequential nucleotides, and each possible combination of three nucleotides encodes one amino acid (with some built-in redundancy in the code and a few exceptions such as the codon for "stop"). In the protein language, groups of amino acid "letters" also form "words", which give rise to the fundamental units of protein secondary structure, namely alpha helices and beta strands, which we will discuss below. Finally, in the DNA language, sentences are formed by collections of words (i.e. collections of codons) and correspond to genes, where we note that a given gene codes for a corresponding protein. An example of a complete sentence in the protein language is a fully folded, compact, biologically active enzyme that is capable of catalyzing a specific biochemical reaction within a living cell. Much of the rest of the book will be built around developing models of a host of biological processes in cells that appeal to the

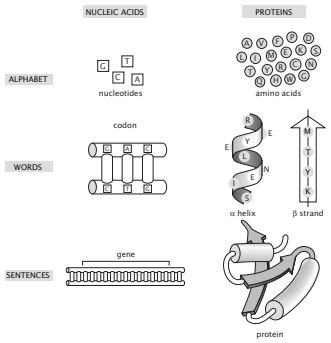

Figure 1.2: Illustration of Crick's "two great polymer languages". The left column shows how nucleic acids can be thought of in terms of letters (nucleotides), words (codons) and sentences (genes). The right column illustrates a similar idea for proteins, where the letters correspond to amino acids, the words to elements of secondary structure such as alpha helices and beta strands, and the sentences to fully folded functional proteins.

macromolecular underpinnings described here.

The sequences associated with nucleic acids and proteins are linked mechanistically through the ribosome which takes nucleic acid sequences (in the form of mRNA) and converts them into amino acid sequences (in the form of proteins). These two polymer languages are linked informationally in the form of the genetic code. In particular, each three letter collection from the nucleic acid language has a corresponding letter that it codes for in the protein language. The determination of this genetic code is one of the great chapters in the history of molecular biology and the results of that quest are shown in fig. 1.4 which shows the information content recognized by the universal translating machine (the ribosome) as it converts the message contained in mRNA into proteins.

1.3 Model Building in Biology

1.3.1 Models as Idealizations

In trying to understand the nature of life, the approach that we take in this book is the same as that followed by scientists for generations. We will approach each system with the goal of gaining some useful insight into its behavior by abstracting and simplifying the highly complex materials of living organisms so that we can apply simple, analytical models that have some predictive value. By definition, we cannot retain a complete atomic-description of each macromolecule. Instead, we aim to select only the relevant properties of the macromolecule that speak to the particular aspect of that molecule's behavior that we are trying

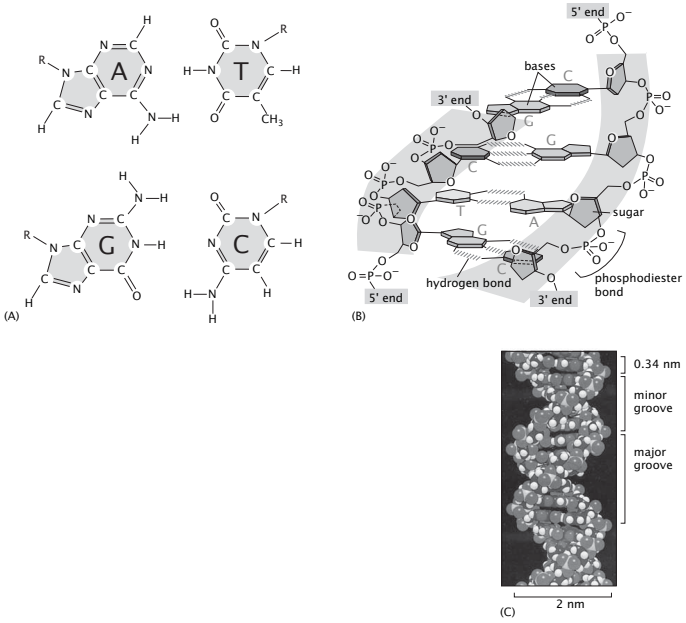

Figure 1.3: The chemical structure of nucleotides and DNA. (A) In DNA, the four distinct subunits or bases are abbreviated A (adenine), T (thymine), G (guanine), and C (cytosine). In these diagrams, carbon is represented by C, oxygen by O, nitrogen by N and hydrogen by H. The letter R indicates attachment to a larger chemical group (the rest of the molecule); for nucleotides, the R group consists of the pentose sugar deoxyribose attached to phosphate. A single line connecting two atoms indicates a single covalent bond and a double line indicates a double covalent bond. The two large bases, A and G, are called purines and the two small bases, C and T, are called pyrimidines. (B) Illustration of how bases are assembled to form DNA, a double helix with two "backbones" made of the deoxyribose and phosphate groups. The four bases are able to form stable hydrogen bonds uniquely with one partner such that A pairs only with T and G pairs only with C. The structural complementarity of the bases enables the faithful copying of the nucleotide sequence when DNA is replicated or when RNA is transcribed. (C) Space-filling atomic model approximating the structure of DNA. The spacing between neighboring base pairs is 0.34 nm.

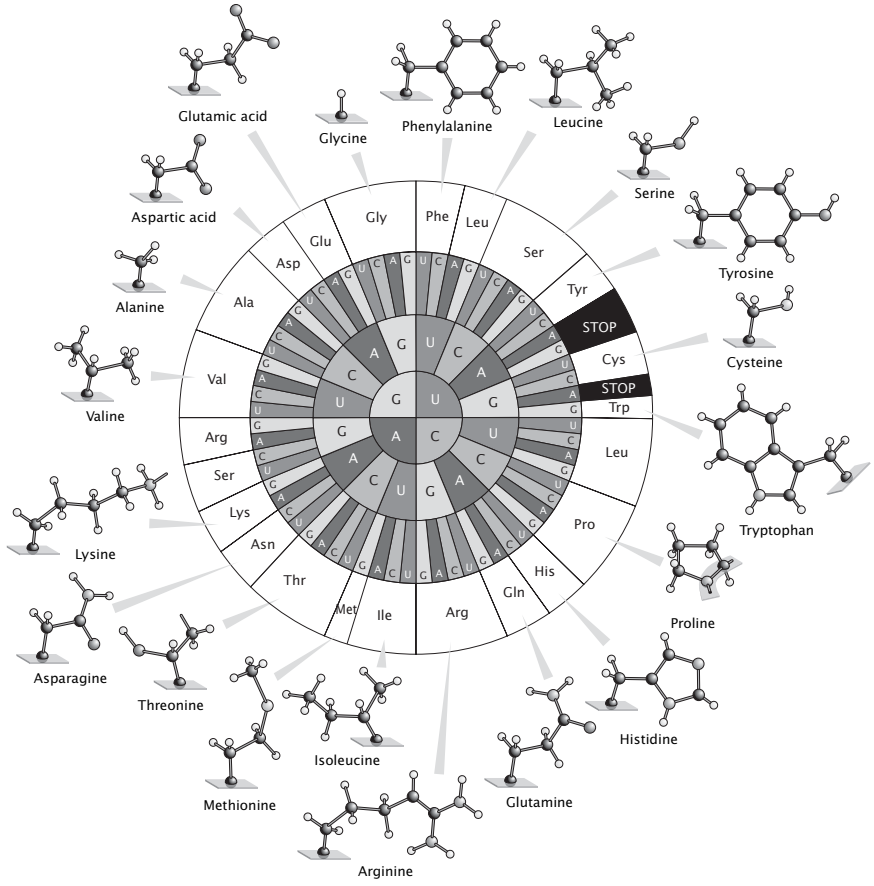

Figure 1.4: Genetic code. In this schematic representation, the first nucleotide in a coding triplet is shown at the center of the ring, the second nucleotide in the middle and the third nucleotide in the outer ring. The amino acids corresponding to each group of triplets are illustrated with their names and atomic structures. Two amino acids, tryptophan and methionine, are encoded by only a single triplet, whereas others including serine, leucine and arginine are encoded by up to six. Three codons do not code for any amino acid and are recognized as stop signals. The unique codon for methionine, AUG, is always used to initiate protein synthesis.

to address. Therefore, it would be inaccurate to say that we have in mind "a" simple model for a complex macromolecule such as DNA. Instead, we will use a suite of simple models that can be thought of as projections of the complex, multifaceted reality of the DNA molecule into a simplified conceptual space. In this section, we will summarize the kinds of physical properties of biological materials that we may wish to consider in isolation. In section 1.4, we will list the kinds of fundamental physical models that will be applied throughout the book.

Biological Stuff Can Be Idealized Using Many Different Physical Models

Often, the essence of building a useful and enlightening model is figuring out what features of the problem can be set aside or ignored. Throughout the book, we will come to view molecules and cells from many different angles. An example of this diversity of outlook is shown in fig. 1.5. The nature of the simplified space, and therefore the nature of the projected DNA model, depends upon the specific question that we are asking. Fig. 1.5 shows five different projections for the meaning of DNA that we will use repeatedly throughout the book.

DNA is one of the most important molecules, both because it is the carrier of genetic information and because of its iconic role in molecular biology. There is a polarity to each strand of this helical polymeric molecule; that is, the two ends of the strand are structurally and chemically distinct, and are called the 5' and 3' ends. In addition, there is heterogeneity along their length, since the chemical identity of the bases from one nucleotide to the next can be different. It is often the informational content of the DNA molecule that garners attention. Indeed, the now-routine sequencing of whole genomes of organisms ranging from bacteria to humans permits us to forget the molecular details of the DNA molecule and to focus just on the sequence of As, Ts, Gs and Cs that determine its information content. This representation will be most useful, for example, when we are trying to discern the history of mutational events that has given rise to diversification of species or when we are trying to calculate the information content of genomes. The second representation abstracts variations in DNA sequence in large blocks where one particular dark gray block represents a portion of sequence that has particularly strong affinity for a protein binding partner, such as the RNA polymerase enzyme that copies the sequence encoded by the DNA into a molecule of messenger RNA. Both these first two representations draw attention to dissimilarities in detailed chemical structure within the DNA molecule. However, sometimes we will consider DNA simply as a physical entity within the cell where the specific nature of its sequence is not relevant. Instead, other properties of the molecule will be emphasized, such as its net distribution of charge, bending elasticity or behavior as a long polymer chain subject to thermal undulations.

Proteins are subject to many different idealizations as well. Just as nucleic acids can be thought of as sequences of nucleotides, we can think of proteins in terms of the linear sequence of amino acids that make them up. The second great polymer language gives rise to proteins that are considered to provide

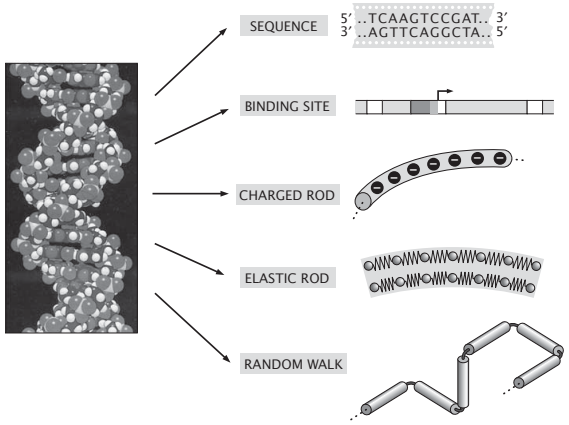

Figure 1.5: Idealizations of DNA. DNA can be thought of as a sequence of base pairs, as a series of binding sites, as a charged rod, as an elastic rod or as a freely jointed polymer arranged in a random walk, depending upon the problem of interest.

most of the important biological functions in a cell, including the ability to catalyze specific chemical reactions. The detailed atomic structure of proteins is substantially more complex than DNA, both because there are twenty rather than four fundamental subunits and because the amino acid chain can form a wider variety of compact three-dimensional structures than can the DNA double helix. Even so, we will extract abstract representations of proteins emphasizing only one feature or another as shown in fig. 1.6 in much the same way we did for DNA. A protein can be represented as a simple linear sequence of amino acids. In cases where this representation contains too much detail, we will elect to simplify further by grouping together all hydrophobic amino acids with a common designation H and all polar amino acids with a common designation P. In these simple models, the physical differences between hydrophobic and polar amino acids (i.e., those that are like oil vs. those that are like water) affect the nature of their interactions with water molecules as well as with each other and thereby dictate the ways that the amino acid chain can fold up to make the functional protein. Similarly, we can think of the three-dimensional folded structure of proteins at different levels of abstraction depending on the question. A common simplification for protein structure is to imagine the protein as a series of cylinders (alpha helices) and ribbons (beta strands) that are connected to one another in a defined way. This representation ignores much of the chemical complexity and atomic-level structure of the protein but is useful for considering the basic structural elements. An even more extreme representation that we will

find to be useful for thinking about protein folding is a class of lattice models of compact polymers in which the amino acids are only permitted to occupy a regular array of positions in space. Fundamentally, the amino acid sequence and folded structure of the protein serve a biological purpose expressed in protein function or activity, and we will also develop classes of models that refer only to the activity of the molecule. For example, the vast majority of proteins in the cell are capable of forming specific binding interactions with one another or with other biomolecules; these binding events are usually important to the protein's biological function. When modeling these kinds of interactions, we will envision proteins as "receptors" that carry binding sites which may be occupied or unoccupied by a binding partner, called a "ligand", and our description will make no further reference to the protein's complex, internal structure. For some models we will focus on changes in activity of proteins considering them as simple two-state systems that can interconvert between two different functional forms. For example, an enzyme that catalyzes a particular biochemical reaction may exist in an "active" or "inactive" state, depending on the presence of ligands or other modifications to its structure.

By performing these simplifications and abstractions we are not intending to deny the extraordinary complexity of these molecules and are fully cognizant that amino acid sequence, folding and compaction properties and conformational changes between states dictate binding affinity, and it is an artificial simplification to consider these properties in isolation from one another. Nonetheless, it is generally useful to consider one single level of description at a time when trying to gain intuition from simple estimates and models, which will serve as our primary emphasis.

We will extend this approach beyond individual macromolecules to large scale assemblies of macromolecules such as, for example, cell membranes made up of millions of individual lipid and protein molecules. Some of the different ways in which we will think about biological membranes are shown in fig. 1.7. Just as with individual macromolecules, these large scale assemblies can be usefully characterized by focusing on one aspect of their physical behavior at a time. For example, membranes can be thought of as bendable, springy elastic sheets, as random surfaces, and even as arrays of electrical elements such as resistors and capacitors. In their interactions with other molecules, membranes often serve as selective barriers that will enable some molecular species to cross while blocking others. Depending upon the particular question being asked, we will exploit different idealizations of the membrane.

Audaciously, we will bring these same idealization techniques to bear even on living cells as shown in fig. 1.8. Although a living cell is exceedingly complex, we will nonetheless choose to extract one physical property of the cell at a time so that we can gain insight from our simple approach. For example, the bacterium *Escherichia coli*, a beloved experimental organism, can for some purposes be thought of as an object carrying an array of protein receptors on its surface. For other purposes, for example when we think of how a bacterium swims through water, we will think of it as an elastohydrodynamic object, that is, a mechanical object that can bend and can interact with the flow of water. The physical

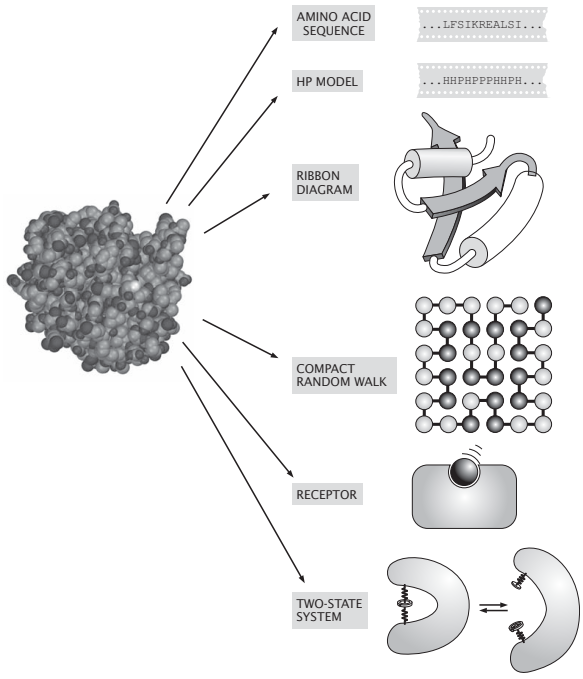

Figure 1.6: Idealizations of protein. Proteins can be thought of as a particular sequence of amino acids, as a simplified sequence reporting only the hydrophobic (H, oil-like) or polar (P, water-like) chemical character of the amino acids, as a collection of connected ribbons and cylinders, as a compact regular polymer on a lattice, as a binding platform for ligands, or as a two-state system capable of interconverting between different functional forms.

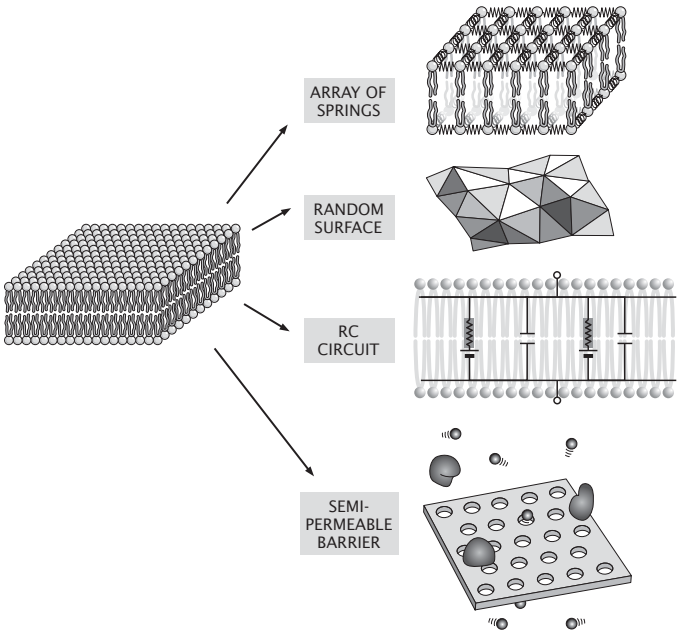

Figure 1.7: Idealization of membranes. A membrane can be modeled as an elastic object which deforms in response to force, as a random surface fluctuating as a result of collisions with the molecules in the surrounding medium, as an electrical circuit element, and as a barrier with selective permeability.

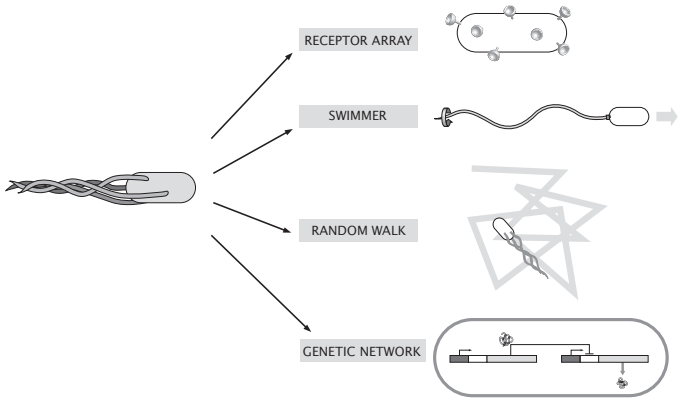

Figure 1.8: *E. coli* idealization. The cell can be modeled as an array of receptors for a ligand of interest, as an elastohydrodynamic object, as a biased random walker moving through water, and as an information processing device characterized by a series of genetic networks.

properties of the path it follows through water can be further abstracted by considering the cell's large-scale motion in the context of a biased random walk, ignoring the hydrodynamic details that actually enable that motion. Finally, because much biological research focuses on how cells alter the expression of their polymer languages in response to changing conditions, we will frequently find it useful to consider the cell as an information processing device, for example as a network governing the flow of information in the form of gene expression. These descriptions are not mutually exclusive, nor are any of them comprehensive. They are again deliberate abstractions made for practical reasons.

Although so far we have emphasized the utility of abstract projections centering on a single physical property for descriptions of biological entities, the same approach is actually extremely useful in consideration of other systems such as solutions of charged ions in water. Various abstractions of the watery medium of life are shown in fig. 1.9. Sometimes we will pretend that a solution is a regular lattice. For example, if we are interested in ligands in solution and their tendency to bind to a receptor, we adopt a picture of the solution in which the ligands are only permitted to occupy specific discrete positions. Although this is an extreme approximation, it is nonetheless immensely valuable for calculations involving chemical equilibrium and remarkably gives accurate quantitative predictions. As the watery medium interacts with the living creatures it contains, sometimes its hydrodynamic flow properties (or viscous properties) are of most interest. We will also see that rates and dynamics play a crucial role. When considering rates of chemical reactions taking place in water we will find it useful to think of water as the seat of diffusive fluctuations. The macromolecules that exist in the watery medium of cells also interact with specific properties of water. For example, some of the chemical groups in proteins

are polar or "hydrophilic" (water-loving), meaning that they are able to form hydrogen bonds with water, while other groups are oil-like and repel water. In the folded protein, the oil-like portions tend to cluster on the inside. Charged molecules can interact with the dielectric character of water molecules, which are neutral overall but do exhibit a slight charge separation.

Each of these representations will resurface multiple times throughout the book in different contexts related to different biological questions. Although none of them can give a complete understanding of the behavior of living cells, each of them serves to provide quantitative insight into some aspect of life and taken as a whole they can begin to project a more realistic image of cells than any single abstraction can do alone.

1.3.2 Cartoons and Models

Biological Cartoons Select Those Features of the Problem Thought to Be Essential

We have argued that the art of model building ultimately reflects a tasteful separation of that which is essential for understanding a given phenomenon from that which is not. As is true of all sciences, the history of experimental progress in biology has always involved researchers constructing explicit conceptual models to help them make sense of their data. In some cases, the theoretical framework developed for biological systems is of precisely the mathematical character that will be presented throughout the book such as illustrated in the use of thermodynamics and statistical mechanics to understand biochemical reactions or in the development of the theory of the action potential. More commonly, much of the important modeling of biological systems has been in the form of visual schematics which illustrate the most essential features in a given biological process and how they interact. In fact, the act of drawing the type of cartoon found in molecular biology textbooks or research papers reflects an important type of conceptual model building and either implicitly or explicitly reflects choices about which features of the problem are really important and which can be ignored.

As an example of the model building that takes place in constructing biological cartoons, consider the structure of one of the most beautiful and intriguing eukaryotic organelles, namely, the mitochondrion. Fig. 1.10 shows a planar section of a mitochondrion as obtained using electron microscopy. Next to the image are two cartoons that describe the structure of this important organelle. This pair of drawings emphasizes the central importance of the membranes which bound the organelle and which separate the interior into different, relatively isolated compartments. Though the three-dimensional model represents more detail regarding the precise geometry of the membranes, the essential *conceptual* elements are the same, namely, a) the mitochondria are closed, membrane-bound organelles, b) the inner membrane is decorated with a series of protrusions which segregate different regions of the mitochondria. Furthermore, these protrusions greatly increase the total amount of surface area of the

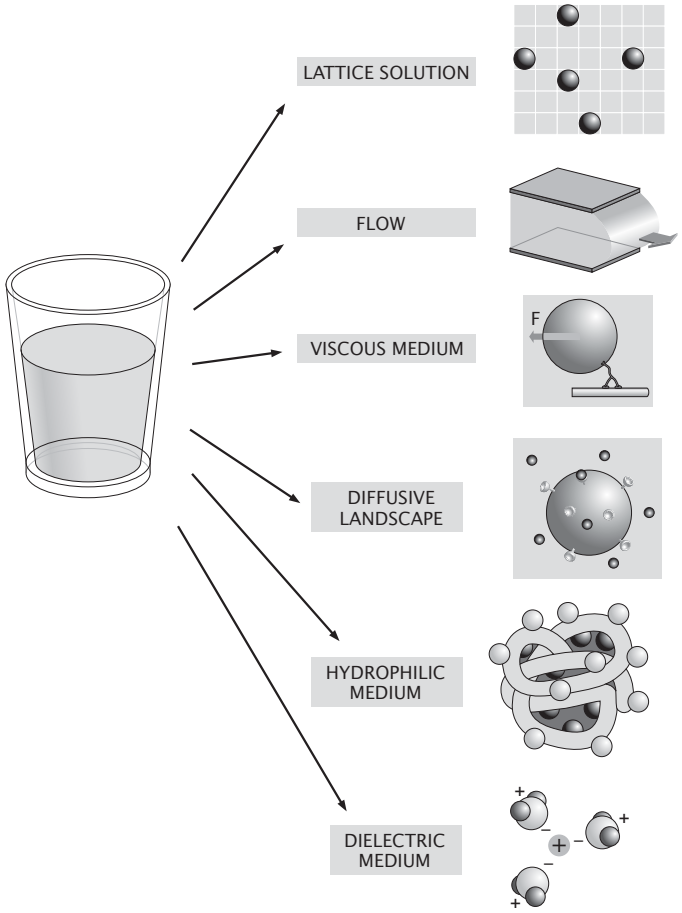

Figure 1.9: Idealization of a solution. A solution can be represented as a lattice of discrete positions where solutes might exist, as a fluid with a mean flow, as a viscous medium that exerts a drag force on objects moving through it, as a fluctuating environment which induces random motion of macromolecules and ligands, as a hydrophilic medium that readily dissolves polar molecules but repels oil-like molecules, and as a dielectric medium which is a poor electrical conductor but can support electrostatic fields.

mitochondrial inner membrane and therefore increase the amount of real estate available for the ATP synthesizing machinery.

The other cartoons in this figure also represent the mitochondrion, but in very different ways, focusing on specific conceptual elements involved in different aspects of mitochondrial function and behavior. The central diagram shows a highly magnified segment of the mitochondrial inner membrane, illustrating the proteins that cross the membrane which are involved in the generation of ATP. A central function of mitochondria in eukaryotic cells is to convert one form of chemical energy, high energy electrons in the compound NADH, into a different form of chemical energy, a phosphodiester bond in ATP (these different forms of chemical energy and their significance will be discussed in more detail in chap. 5). The process of performing this conversion requires the cooperative action of several different large protein complexes that span the inner mitochondrial membrane. Sequential transfer of electrons from one protein complex to another results in the transport of hydrogen ions from the mitochondrial matrix across the inner membrane into the intermembrane space, establishing a gradient with an excess of hydrogen ions on the outside. In the final step, the hydrogen ions are allowed to travel back down their concentration gradient into the mitochondrial matrix, by passing through a remarkable protein machine called ATP synthase, which we will revisit several times throughout the rest of the book. ATP synthase catalyzes the construction of a molecule of ATP by combining a molecule of ADP and an inorganic phosphate ion. The precise molecular details of this process are fascinating and complex; our intention here is merely to point out that this kind of cartoon represents an abstraction of the mitochondrion that emphasizes a completely different set of components and concepts than the membrane-focused cartoons shown above. The final cartoon illustrates yet another aspect of mitochondrial life; these organelles contain their very own DNA genomes, completely distinct from the genome found in the cell's nucleus. As indicated in the diagram, the genome is a circular piece of DNA (like the genome found in bacterial cells, but very different from eukaryotic chromosomes which are linear pieces of DNA). Each mitochondrion contains several copies of the genome, and they are distributed between the two daughters when the mitochondrion divides.

Each of these different kinds of cartoons are valid and informative models of certain aspects of the mitochondrion, but they serve essentially non-overlapping conceptual purposes. Furthermore, each cartoon represents the culmination of thousands of separate experiments covering decades of hard-won, detailed knowledge; experts have carefully sifted the raw data and made difficult decisions about which details are important and which are dispensable in considering each of these aspects of mitochondrial reality. Throughout the rest of this book, we will take advantage of the hard work of conceptualization that has already gone into constructing these kinds of diagrams. Our usual goal will be to take these existing biological models one step further, by rendering them into a mathematical form.

Quantitative Models Can Be Built by Mathematicizing the Cartoons

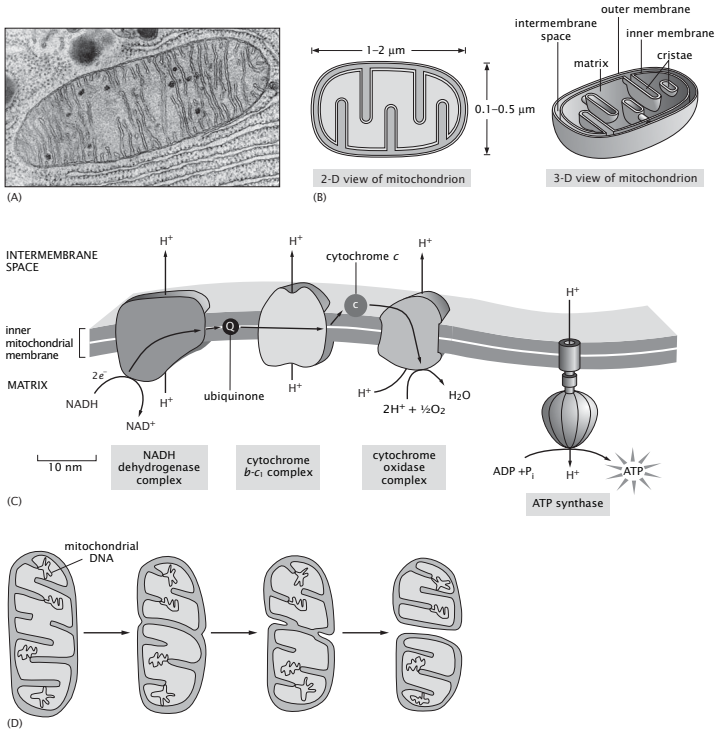

Figure 1.10: Several different ways of illustrating mitochondria. (A) Thin-section electron micrograph showing a mitochondrion found in a cell within the pancreas of a bat. (B) Diagrams showing the arrangement of membranes dividing the mitochondrion into distinct compartments. While the outer membrane forms a smooth capsule, the inner membrane is convoluted to form a series of cristae. Distinct sets of proteins are found in the matrix (inside of the inner membrane) and in the intermembrane space (between the inner and outer membranes). (C) Schematic illustration of the major proteins involved in the electron transport chain and in ATP synthesis in the inner membrane of the mitochondrion. The overall purpose of the complex chemical reactions carried out by this series of proteins is to catalyze the creation of the important energy carrier molecule ATP. (D) Illustration of the distribution and partitioning of mitochondrial DNA during the process of mitochondrial division (fission). (A, from D. W. Fawcett, The Cell: An Atlas of Fine Structure, Philadelphia: W. B. Saunders & Co., 1966; C, from B. Alberts, *et al.*, Molecular Biology of the Cell, 4th ed. New York: Garland Science, 2002.)

One of the key mantras of the book is that the emergence of quantitative data in biology requires that the cartoons of molecular biology be mathematicized. In particular, in some cases only by constructing a mathematical model of a given biological problem can the model be put on the same footing as the data itself. A concrete example of this that will arise in chap. 19 is the measurement of gene expression as a function of the distance between two binding sites on DNA as shown in fig. 1.11. The mechanistic basis of this data is that a certain protein (Lac repressor) binds simultaneously at two sites on the DNA and forms a loop of the DNA between the two binding sites. When the Lac repressor protein is bound to the DNA, it prevents the binding of the protein machine that copies the DNA into RNA. The strength of this effect (repression) can be measured quantitatively. The experimental data shows that the amount of repression depends in a complex way on the precise distance between the two binding sites.

Our argument is that a cartoon-level model of this process, while instructive, provides no quantitative basis for responding to this data. For example, why are there clear periodic peaks in the data? Why does the height of the peaks first increase and then decrease? In chap. 19, we will derive a mathematical model for this process using the framework of statistical mechanics that accurately predicts the shapes and sizes of these features in the data while at the same time highlighting some features of DNA mechanics that remain unclear. Further, by constructing a physical model in mathematical terms of this process, new experiments are suggested which sharpen our understanding of transcriptional regulation. In nearly every chapter, we will repeat this same basic motif: some intriguing quantitative observation on a biological system can be described first in terms of a cartoon, which can then be recast in mathematical form. Once the mathematical version of the model is in hand, we use the model to examine previous data and to suggest new experiments.

As a result of examples like this (and many others to be seen later) one of the central arguments of the present chapter is that there are many cartoons in molecular biology and biochemistry which have served as conceptual models of a wide variety of phenomena and that reflect detailed understanding of these systems. To make these models quantitatively predictive they need to be cast in mathematical form, this in keeping with our battle cry that quantitative data demands quantitative models. As a result, one way of viewing the chapters that follow is as an attempt to show how the cartoons of molecular biology and biochemistry have been and can be "translated into a mathematical form".

1.4 Quantitative Models and the Power of Idealization

Quantitative models of the natural world are at the very heart of physics and in the remainder of the book we will illustrate how pervasive these ideas have become in biology as well. This approach is characterized by a rich interplay

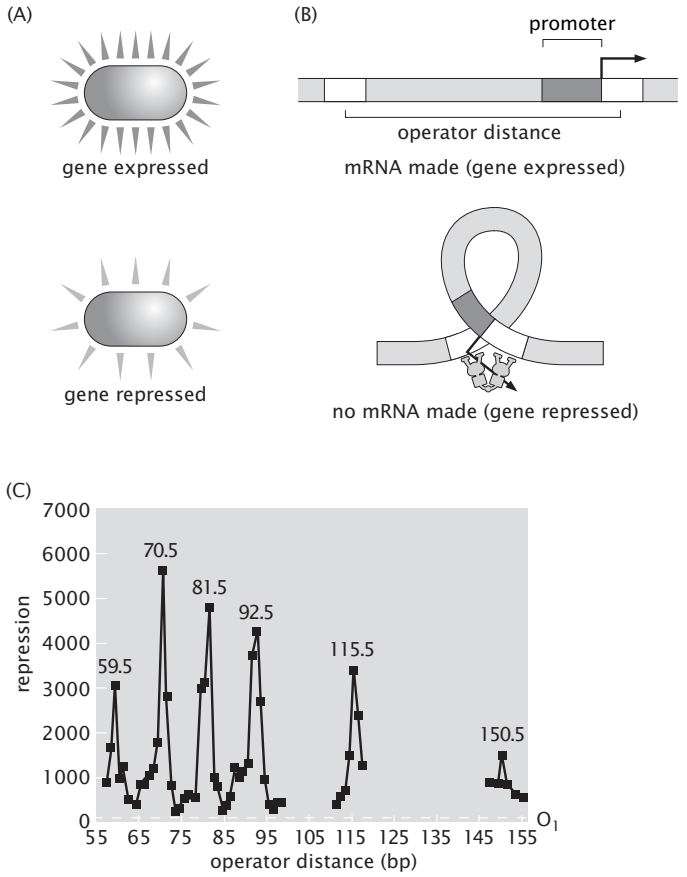

Figure 1.11: Qualitative and quantitative illustrations of gene repression. (A) Bacteria containing a gene whose expression level is regulated by the Lac repressor protein may exist in different states. When repression is low, the gene is expressed, and large amounts of the gene product are present. If the gene being regulated encodes a fluorescent protein, the bacteria will glow brightly when the gene is expressed. In contrast, when the gene is repressed, relatively little fluorescent protein will be produced, and the bacteria will glow dimly. (B) Schematic showing events at the level of the DNA during gene expression and repression. During repression, the Lac repressor protein binds to two distinct operator sites on the DNA (white boxes) and forces the DNA to form a loop. This prevents binding of the protein machine that copies the DNA into mRNA. When no mRNA is produced, the gene is repressed. (C) Quantitative measurements of the strength of gene repression as a function of the exact distance (expressed in base pairs) between the two operator sites. (C, data from J. Müller *et al.*, *J. Mol. Biol.*, 257:21, 1996.)

between theory and experiment in which the consequences of a given model are explored experimentally. In return, novel experimental results beckon for the development of new theories. To illustrate the way in which particular key models get recycled again and again in surprisingly varied contexts, we consider one of the most fundamental physical models in all of science, namely, the simple spring (also known as the harmonic oscillator).

1.4.1 On the Springiness of Stuff

The concept of springiness arises from the confluence of a very important mathematical idea of a Taylor series (explained in detail in "Math Behind the Models" on pg. 273) and an allied physical idea known as Hooke's law (the same Hooke that ushered in microscopy in biology). These ideas will be developed in detail in chap. 5 and for now we content ourselves with the conceptual framework. The fundamental mathematical idea shared by all "springs" is that the potential energy for almost any system for small displacements from equilibrium is well approximated by a quadratic function of the displacement. Mathematically, we can write this as

$$\text{Energy} = \frac{1}{2}kx^2. \tag{1.1}$$

What this equation states is that the potential energy increases as the square of the displacement x away from the equilibrium position. The "stiffness" k is a measure of how costly it is to move away from equilibrium and reflects the material properties of the spring itself.

Another way of thinking about springs is that they are characterized by a restoring force that is proportional to how far the spring has been displaced from its equilibrium position. Mathematically, this idea is embodied in the equation

$$F = -kx, \tag{1.2}$$

where F is the restoring force, x is the displacement of the spring from its equilibrium position and k is the so-called spring constant (the stiffness). The minus sign signals that the force is a restoring one towards $x = 0$ which designates the equilibrium position. This result is known as Hooke's law. When stated in this way, this kind of physical model gives the impression of an abstract example of masses on frictionless tables with pulleys and springs. But empirically, scientists have found that precisely this mathematical model arises naturally in many different practical contexts, as shown in fig. 1.12.

In biology, the simple spring comes cloaked in many different disguises as we will see throughout the book. From a technological perspective, many of the key single-molecule techniques used to study macromolecules and their assemblies invoke this description. Both optical tweezers and the atomic force microscope can be mapped precisely and unequivocally onto spring problems. However, as shown in the figure, there are many other unexpected examples in which the problem can be recast in a way that is mathematically equivalent to the simple spring problem introduced above. Simple spring models have been invoked when

considering the bending of DNA, the beating of the flagellum on a swimming sperm and even the fluctuations of membranes at the cell surface. As these are all mechanical processes, it is fairly straightforward to see how the idea of a spring can apply to them. But the real power of the simple spring model will be revealed when we see how it can be directly applied to non-mechanical problems. For example, when we discuss biochemical reactions in a cell, we will imagine molecules moving on an "energy landscape" where they accumulate in "potential wells". The mathematics of simple springs will help us understand the rates that govern these biochemical reactions. The ideas of harmonic oscillators and potential wells will also illuminate our exploration of protein conformational changes. Even more abstractly, we will see that changes in gene expression over time for certain interesting kinds of genetic networks can also be surprisingly well-described using the mathematics of simple harmonic oscillators.

1.4.2 The Toolbox of Fundamental Physical Models

The overall strategy of this book will be to apply this kind of fundamental physical thinking to the widest possible range of biological problems. Through careful analysis of the quantitative assumptions that go into the analysis of biological data, it has become clear that a mere handful of fundamental physical models are often sufficient to provide a rigorous framework for interpretation of many kinds of quantitative biological data.

Beyond the harmonic oscillator introduced above, there are between five and ten other key physical concepts that require mastery. A scientist who understands and appreciates the applications of these concepts in biology is afforded a powerful framework for zeroing in on the most important, unexplained open questions. It is not our contention that this family of broad reaching physical models is sufficient to explain biological phenomena, far from it. Rather, we contend that any observations that can be completely quantitatively accounted for within the framework of one of these models require no further speculation of unidentified mechanism, while data that cannot be embraced within this physical framework cry out for more investigation and exploration.

Our choice of this set of concepts is purely utilitarian. We are not attempting to write a book surveying all of physics, nor surveying all of biology. Modern, cutting-edge research at the interdisciplinary interface between biology and physics demands that scientists trained in one discipline develop a working knowledge of the other. Students currently training in these fields have the opportunity to explore both in detail and understand their complementarity and interconnectedness. In this book, we have attempted to address all three audiences, physicists curious about biology, biologists curious about physics and broad-minded learners ready to delve into both. Our hope is that these ideas will provide a functional foundation of some of the key physical concepts most directly relevant to biological inquiry. At the same time, we hope to provide enough examples of specific biological phenomena to allow a physical scientist to identify further, important unsolved problems in biology.

The bulk of this book comprises detailed dissections and case studies of these

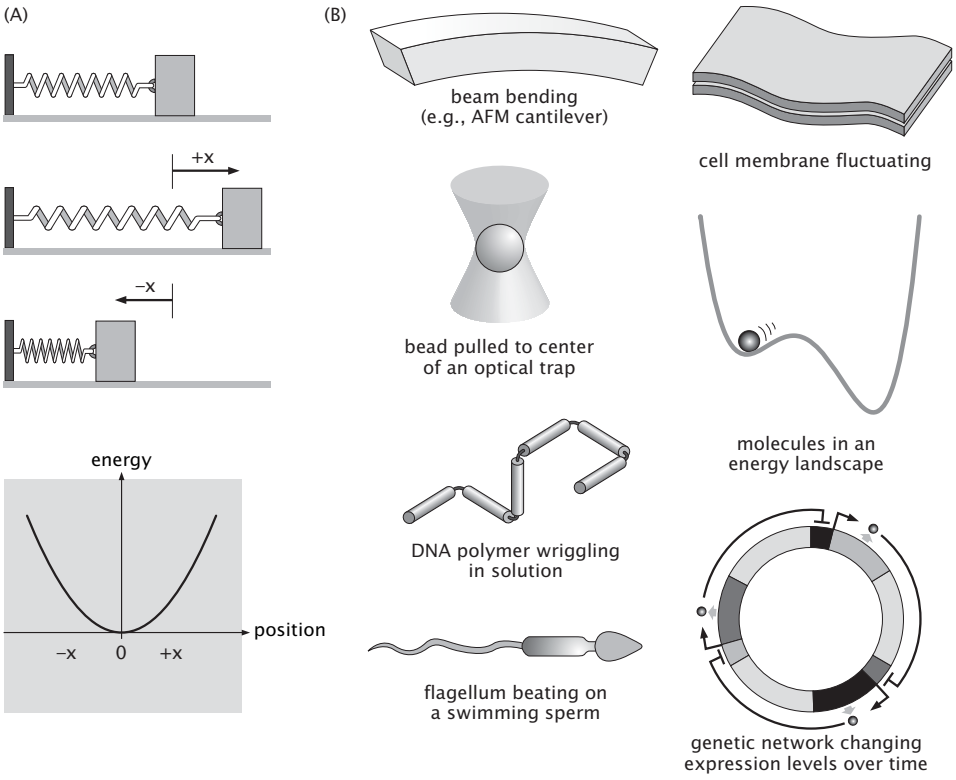

Figure 1.12: Gallery showing some of the different ways in which the idea of a spring is used in biological model building. (A) Classical physics representation of the harmonic oscillator. A simple spring is attached at one end to an immovable wall, and at the other end to a block-shaped mass that can slide on a frictionless table. The position of the mass changes as the spring extends and compresses. The graph at the bottom shows the energy of the system as a function of the mass position. When the spring is relaxed at its resting length, the energy is at a minimum. (B) Multiple manifestations of the simple harmonic oscillator relevant to biological systems.

fundamental concepts and concludes with several chapters showing applications of several of these concepts simultaneously. As noted above, many individual chapters each focus on a single class of physical model. The key broad physical foundations are:

- *Simple Harmonic Oscillator.* (chap. 5)
- *Ideal Gas and Ideal Solution Models.* (chap. 6)
- *Two-Level Systems and the Ising Model.* (chap. 7)
- *Random Walks, Entropy and Macromolecular Structure.* (chap. 8)
- *Poisson-Boltzmann Model of Charges in Solution.* (chap. 9)
- *Elastic Theory of One-Dimensional Rods and Two-Dimensional Sheets.* (chaps. 10 and 11)
- *Newtonian Fluid Model and the Navier-Stokes Equation.* (chap. 12)
- *Diffusion and Random Walks.* (chap. 13)
- *Rate Equation Models of Chemical Kinetics.* (chap. 15)

1.4.3 The Role of Estimates

In this era of high speed computation, the temptation to confront new problems in their full mathematical complexity is seductive. Nevertheless, the philosophy of the present book advocates a spirit that is embodied in a story, perhaps apocryphal, of the demise of Archimedes. Plutarch tells us that Archimedes was deeply engaged in a calculation, which he was performing with a stick in the sand, at the time when the Roman army invaded Syracuse. When confronted by a Roman soldier and asked to identify himself, Archimedes told the soldier to wait until he was done with his calculation. The soldier replied by impaling Archimedes, bringing one of the great scientific minds of history to an unjust and early end. Our reason for recounting this story is as a reminder of the mathematical character of many of the models and estimates that we will report on in this book (but hopefully with less dire personal consequences). In particular, we like to imagine that a model is of the proper level of sophistication to help build intuition and insight if, indeed, its consequences can be queried by calculations carried out with a stick in the sand, without recourse to sophisticated computers or detailed references.

Throughout the book we will repeatedly make use of estimates to develop a feeling for the numbers associated with biological structures and biological processes. This philosophy follows directly from the biological tradition articulated by the pioneering geneticist Barbara McClintock who emphasized the necessity of developing a "feeling for the organism", in her case the crop plant maize. We believe that a quantitative intuition is a critical component of having a feeling for the organism. Indeed, we think of model building as proceeding through a

series of increasingly sophisticated steps starting with simple estimates to develop a feel for the relevant magnitudes, followed by the development of toy models that can (usually) be solved analytically, followed in turn by models with an even higher degree of realism that might require numerical resolution. The estimates are all meant to be sufficiently simple that they can be done on the back of an envelope or written with a stick in the sand on a desert island. In keeping with this purpose, we will rarely use more than one significant digit in our estimates. Naturally, this introduces somewhat less precision, for example, an estimate may yield an answer of 100 molecules in a system when the precisely measured value is actually 42, but has the benefit of greater simplicity and clarity. In general we will consider an estimate to be successful if the result is within an order of magnitude of the measured value. In cases where our estimates deviate more than this, it will be clear that we have made an incorrect assumption and we will use this as an opportunity to learn something more about the problem of interest.

The precept that animates much of the book is that we are closest to understanding a phenomenon, not when we unleash the full weight of a powerful numerical simulation of the problem at hand, but rather when engaged in developing a feel for the scaling of the quantities of interest and their associated numerical values.

Taylor and the Cover of Life Magazine. One of the greatest stories of the power and subtlety of estimation is associated with G. I. Taylor. During the early days of the atomic bomb, the energetic yield of the first explosion was classified as a military secret. This did not keep those who make such decisions from releasing several time lapse photos of the explosion which showed the radius of the fireball as a function of time. On this meager information, shown on the cover of *Life* magazine, Taylor was able to use the principles of fluid mechanics to *estimate* the yield of the explosion, an estimate that was exceedingly close to the secret value of the actual yield.

Interestingly, this famous episode from the cultural history of the physical sciences could have happened just as well in biology. Pictures like that shown in fig. 1.13 were one of the thrilling outcomes of the early days of the use of electron microscopy in biology. Here too, by observing the physical region occupied by the DNA molecule from the ruptured bacterial virus (bacteriophage), it is possible to make a highly simplified estimate of the *genomic* size (i.e. the length in basepairs) of the bacteriophage.

Why is it useful to know the number of copies of a given molecule that is controlling the expression of a gene of interest within an *E. coli* cell? What merit is there in estimating the rate of actin polymerization at the leading edge of a motile cell? What is the value in knowing the time it takes a given protein to diffuse a millimeter? First, simple estimates serve as a reality check to help us see whether or not our impressions of how a system works are reasonable at all. Second, a sense of the numbers can tell us what kinds of physical constraints are in play. For example, by knowing the size of a bacterium and its swimming speed, we can construct a characteristic number (the Reynolds number) that tells us about the importance of acceleration in the dynamical problem of

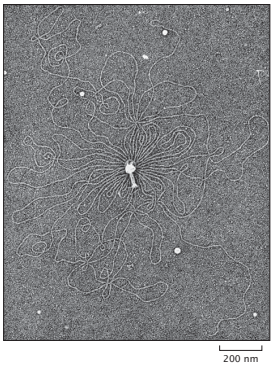

Figure 1.13: Electron microscopy image of a bacteriophage genome which has escaped its capsid. Simple arguments from polymer physics can be used to estimate the genomic size of the DNA by examining the physical size of the randomly spread DNA. We will perform these kinds of calculations in chap. 8. (Adapted from G. Stent, Molecular Biology of Bacterial Viruses, San Francisco: W. H. Freeman, 1963.)

swimming. Similarly, a knowledge of the number of the specific DNA-binding proteins known as transcription factors that control the expression of a gene of interest can tell us about the importance of fluctuations in the gene expression process.

1.4.4 On Being Wrong

When performing simple quantitative estimates, or when attempting to write simple mathematical models to describe some aspect of a biological process, we are easily tempted to worry about whether our estimates and models are "right" or "wrong". There are at least three different ways in which a model might justly be called "wrong", but we believe that committing each of these kinds of wrongness can be a useful and worthwhile experience.

One category of wrong models might fail because they use an inappropriate type of fundamental physical model to abstract an aspect of a particular biological problem. This kind of wrongness might be analogized as asking, how many hours can fit in the back of a pickup truck? The investigator will quickly learn that the predictions of such a model will simply have no relevance to the actual experiment or biological phenomenon being considered. It is inevitable that any person learning to conceptualize the kinds of complicated issues we explore in biology will make some of these kinds of mistakes at first. With time and experience, these errors will become less frequent.

A second category of wrong models might fail because they do not include enough detail about the system. For example, when we try to predict how fast

E. coli bacterial cells can swim through water, we might be able to write an equation that correctly predicts the order of magnitude of the average speed of the cells, but fails to provide any basis for predicting how the speed might fluctuate over time. This kind of error, we believe, is not so much wrongness as incompleteness. As we have argued, it is actually extremely important that physical models be incomplete in order to be simple enough to help in building human intuition. A model that is nearly as complex as the system itself may be able to make very accurate predictions, but will not be useful for promoting understanding.

The third category of wrong models is by far our favorite. These are the wrong models that drive breakthrough experiments. When a person who is well familiar with a particular biological system writes down a quantitative model for that system incorporating the correct basic physical models and uses the correct order-of-magnitude estimates for relevant parameters nevertheless finds that it makes dramatically incorrect predictions, this is almost always a golden opportunity to learn something fundamentally new about the system. Science advances when people notice that something doesn't quite make sense, that something is puzzling, that something seems to be missing. It is our hope that readers of this book will develop sufficient confidence in their ability to create appropriate models that when they find a model seems to be wrong in this way they will seize the opportunity to make a new scientific discovery.

1.4.5 Rules of Thumb: Biology by the Numbers

A mastery of any field requires knowledge of a few key facts. For the physicist that uses quantum mechanics, this means knowing the ground state energy of the hydrogen atom, the charge on an electron, Planck's constant and several other key physical parameters. In physical biology, there are similarly a small set of numbers that are worth carrying around as the basis of simple estimates of the kind one might carry out with a stick in the sand. In that spirit, the table of key numbers given in table 1.4.5 are those that we use and find useful when thinking about physical biology of the cell. In the chapters that follow, these key numbers will permit us to form the habit of making quantitative estimates as a preliminary step when confronting new problems.

Throughout the remainder of the book, we will repeatedly make reference to certain numerical rules of thumb such as that the volume of an *E. coli* cell is roughly 1 μm^3 or that the natural energy scale of physical biology is k_BT (i.e. Boltzmann's constant multiplied by room temperature in degrees Kelvin) $\approx$ 4pN nm (piconewton-nanometers). It is convenient to express this natural energy scale in terms of piconewtons and nanometers when we are considering the forces generated by protein molecular motors moving along DNA, for example, but this natural energy scale can also be expressed in different units depending on the type of calculation we are doing. When considering processes involving charge and electricity, it is convenient to recall that k_BT is approximately equal to 25 meV (millielectron volts). When considering biochemical reactions, it is more useful to express k_BT in biochemical units, as 0.6 kcal/mol (kilocalories

per mole) or 2.5 kJ/mol (kilojoules per mole).

The quantitative rules of thumb tabulated here will serve as the basis of our rough numerical estimates that could be carried out using a stick in the sand without reference to books, papers or tables of data. Where do these numbers come from? Each comes from the results of a long series of experimental measurements of many different kinds. Right now we will simply assert their values, but later in the book we will provide numerous examples showing how these kinds of measurements have been made. How precisely must we know these numbers? As stated earlier, our approach is to be generally satisfied in those cases where our numerical estimates are of the right order of magnitude. To that end, we make certain gross simplifications. For example, when thinking about proteins and the amino acids that make them up, we will often adopt a minimalistic perspective in which the mass of amino acids is *approximated* to be 100 Da (Daltons, or grams per mole). In actuality, real amino acids have masses which range over a wider set of values from 75 to 204 Da. Likewise, the rest of the numbers in this table represent only rough average values to use as a starting point for quantitative estimates. As we progress through the examples in the rest of the book, these values will become increasingly familiar.

Quantity of interest	Symbol	Rule of thumb
E. coli		
Cell volume	$V_{E.coli}$	$\approx 1\ \mu\text{m}^3$
Cell mass	$m_{E.coli}$	$\approx$ 1pg
Cell cycle	$t_{E.coli}$	≈ 3000 s
Cell area	$A_{E.coli}$	$\approx 6\ \mu\text{m}^2$
Genome Length	$N_{bp}^{E.coli}$	$\approx 5 \times 10^6$ bp
Swimming speed	$v_{E.coli}$	$\approx 20\ \mu$m/s
Yeast		
volume of cell	V_{yeast}	$\approx 60\mu\text{m}^3$
Mass of cell	m_{yeast}	≈ 60 pg
diameter of cell	d_{yeast}	$\approx 5\ \mu$m
Cell cycle time	t_{yeast}	≈ 200 min
Genome Length	N_{bp}^{yeast}	$\approx 10^7$ bp
Organelles		
Diameter of nucleus	$d_{nucleus}$	$\approx 5\ \mu$m
Length of mitochondrion	l_{mito}	$\approx 2\ \mu$m
Diameter of transport vesicles	$d_{vesicle}$	≈ 50 nm
Water		
Volume of molecule	V_{H_2O}	$\approx 10^{-2}$ nm^3
Density of water	ρ	1 g/cm^3
Viscosity of water	η	≈ 1 centipoise (10^{-2} g/cm s)
Hydrophobic embedding energy	$\approx E_{hydr}$	25 cal/mol A^2
DNA		
Length per base pair	l_{bp}	$\approx 1/3$ nm
Volume per base pair	V_{bp}	≈ 1 nm^3
charge density	λ_{DNA}	2 e/0.34 nm
Persistence length	ξ_{P}	50 nm
Amino acids and Proteins		
Radius of "Average" Protein	$r_{protein}$	≈ 2 nm
Volume of "Average" Protein	$V_{protein}$	≈ 25 nm^3
Mass of "Average" Amino Acid	M_{aa}	≈ 100 Da
Mass of "Average" Protein	$M_{protein}$	$\approx 30{,}000$ Da
Protein concentration in cytoplasm	$c_{protein}$	≈ 300 mg/ml
Characteristic force of protein motor	F_{motor}	≈ 5 pN
Characteristic speed of protein motor	v_{motor}	≈ 200 nm / s
Diffusion constant of "Average" Protein	$D_{protein}$	$\approx 100\ \mu\text{m}^2$/s
Lipid Bilayers		
Thickness of lipid bilayer	d	≈ 5 nm
Area per molecule	A_{lipid}	$\approx \frac{1}{2}$ nm^2
Mass of lipid molecule	m_{lipid}	≈ 800 Da

1.5 Summary and Conclusions

When applying physical and quantitative thinking to biological problems, two different skills are critically important. First, the biological system must be described in terms of fundamental physical models that can help to shed light on the behavior of the system. Fewer than ten truly basic physical models are able to provide broad explanatory frameworks for many problems of interest in biology. Second, appropriate quantitative estimates to give order-of-magnitude predictions of sizes, numbers, time-scales, and energies involved in a biological process can serve as a reality check that our conception of the system is appropriate and can help predict what kinds of experimental measurements are likely to provide useful information. Armed with these skills, an investigator can begin to systematize the large amounts of detailed quantitative data generated by biological experimentation and determine where new scientific breakthroughs might be sought.

1.6 Further Reading

B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, **Essential Cell Biology**, Garland Publishing Inc., New York: New York, 1998. This book provides a discussion of much of the biological backdrop for our book.

M. Schaechter, J. L. Ingraham and F. C. Neidhardt, **Microbe**, ASM Press, Washington DC, 2006. This book is full of insights into the workings of microscopic organisms.

L. Stryer, **Biochemistry**, W. H. Freeman and Company, New York: New York, 1995. We like to imagine our readers with this book at their side. Stryer is full of interesting insights presented in a clear fashion.

R. Schleif, **Genetics and Molecular Biology**, The Johns Hopkins University Press, Baltimore: Maryland, 1993. Schleif's book discusses many of the same topics found throughout our book from a deeper biological perspective, but with a useful and interesting quantitative spin.

S. Vogel, **Comparative Biomechanics - Life's Physical World**, Princeton University Press, Princeton: New Jersey, 2003. Though largely aimed at describing problems at a different scale than those described in our book, Vogel's book (as well as his many others) present an amazing merger of biology and physical reasoning.

The spirit of physical biology we are interested in developing here has already been described in a number of other books. Some of our favorites are: K. Dill and S. Bromberg, **Molecular Driving Forces**, Garland Press, New York: New

York, 2003. Despite a deceptive absence of difficult mathematics, their book is full of insights, subtlety and tasteful modeling. D. Boal, **Mechanics of the Cell**, Cambridge University Press, Cambridge: England, 2002. Boal's book illustrates the way in which physical reasoning can be brought to bear on problems of biological significance. We have found that many of the exercises at the back of each chapter are the jumping off point for the analysis of fascinating biological problems. J. Howard, **Mechanics of Motor Proteins and the Cytoskeleton**, Sinauer Associates, Inc., Sunderland: Massachusetts, 2001. Howard's book is a treasure trove of information and concepts. P. Nelson, **Biological Physics: Energy, Information, Life**, W. H. Freeman and Company, New York: New York, 2004. This book discusses a wide range of topics at the interface between physics and biology. G. B. Benedek and F. M. H. Villars, **Physics With Illustrative Examples from Medicine and Biology**, Springer-Verlag, Inc., New York: New York, 2000. This three-volume series should be more widely known. It is full of useful insights into everything ranging from the analysis of the Luria-Delbruck experiment to the action potential. K. Sneppen and G. Zocchi, **Physics in Molecular Biology**, Cambridge University Press, Cambridge: United Kingdom, 2005. This book presents a series of vignettes that illustrate how physical and mathematical reasoning can provide insights into biological problems. M. B. Jackson, **Molecular and Cellular Biophysics**, Cambridge University Press, Cambridge: United Kingdom, 2006. Jackson's book covers many themes in common with our own book and is especially strong on ion channels and electrical properties of cells.

J. Harte, **Consider a Spherical Cow**, University Science Books, Sausalito: California, 1988. This book does a beautiful job of showing how order of magnitude estimates can be used to examine complex problems. Those that are intrigued by this book might consider dipping into his **Consider a Cylindrical Cow** as well.

J. Cohen, "Mathematics is biology's next microscope, only better; Biology is Mathematic's next physics, only better". PLoS Biology, **2**, e439 (2004). This article argues for the way that quantitative reasoning will enrich biology and for the ways that biology will enrich mathematics.

1.7 References

D. Bray, "Reasoning for results", Nature, **412**, 863 (2001). Source of Darwin quote at the beginning of the chapter.

D. W. Fawcett, **The Cell, Its Organelles and Inclusions: An Atlas of Fine Structure**, W. B. Saunders and Company, Philadelphia: Pennsylvania, 1966. Fawcett's atlas is full of wonderful electron microscopy images of cells.

J. Müeller, S. Oehler and B. Müller-Hill, "Repression of *lac* promoter as a function of distance, phase and quality of an auxiliary *lac* operator", J. Mol. Biol., **257**, 21 (1996).

G. Stent, **Molecular Biology of Bacterial Viruses**, W. H. Freeman, San Francisco:CA, 1963.